International Journal of Electrical and Computer Engineering (IJECE)
Vol. 13, No. 2, April 2023, pp. 1483~1490
ISSN: 2088-8708, DOI: 10.11591/ijece.v13i2.pp1483-1490

Alpha-divergence two-dimensional nonnegative matrix factorization for biomedical blind source separation

Abd Majid Darsono, Toh Cheng Chuan, Nurulfajar Abd Manap, Nik Mohd Zarifie Hashim
Centre for Telecommunication Research and Innovation, Faculty of Electronic and Computer Engineering,
Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

Article Info

Article history:

Received May 23, 2021
Revised Jul 27, 2022
Accepted Sep 06, 2022

Keywords:

Alpha-divergence
Multiplicative update
Nonnegative matrix factorization
Sparseness constraints

ABSTRACT

An alpha-divergence two-dimensional nonnegative matrix factorization (NMF2D) for biomedical signal separation is presented. NMF2D is a popular approach for retrieving low-rank approximations of nonnegative data such as image pixel, audio signal, data mining, pattern recognition and so on. In this paper, we concentrate on biomedical signal separation by using NMF2D with alpha-divergence family which decomposes a mixture into two-dimensional convolution factor matrices that represent temporal code and the spectral basis. The proposed iterative estimation algorithm (alpha-divergence algorithm) is initialized with random values, and it updated using multiplicative update rules until the values converge. Simulation experiments were carried out by comparing the original and estimated signal in term of signal-to-distortion ratio (SDR). The performances have been evaluated by including and excluding the sparseness constraint which sparseness is favored by penalizing nonzero gains. As a result, the proposed algorithm improved the iteration speed and sparseness constraints produce slight improvement of SDR.

Corresponding Author:

Abd Majid Darsono
Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka (UTeM)
Hang Tuah Jaya, Durian Tunggal 76100 Melaka, Malaysia
Email: abdmajid@utem.edu.my

1. INTRODUCTION

In recent years, nonnegative matrix factorization (NMF) [1], [2] plays an important role in various fields. The objective of NMF is to locate an estimated factorization for a nonnegative matrix V into two nonnegative networks W and H. The columns of W are called basis functions, while the rows of H represent the hidden nonnegative sources that correspond to each basis function. In comparison, NMF is better than other matrix factorization approaches such as principal component analysis (PCA) or independent component analysis (ICA) [3] because nonnegativity constraint in NMF produces sparse results which are advantageous [4], [5]. NMF claims that it can be easily solved by multiplicative update (MU) rules [6]–[16].

A special case that been discussed for years is biomedical signal processing [17]–[19] where there is a need in decomposition on heart and lung sound. This is due to clean and plain lung sound leads to convenient in the diagnosis of lung condition when interference of heart sound is eliminated via NMF approach [20]. In this paper, progression of NMF is stretched out into two-dimensional nonnegative matrix factorization (NMF2D) model [21] so as to give division that can catch the temporal dependency of the frequency designs inside the source proficiently. The 2D of NMF2D refers to two dimensional which is temporal code and spectral basis. The cost function that used to compute the MU rules algorithm is α-divergence. The α-divergence general framework is proposed by Cichocki *et al.* [22] and it will be further discussed in the following section. The performance is evaluated through comparing original and estimated signal in term of SDR. SDR [23] is signal-to-distortion ratio which is the ratio that compare original signal to estimated signal in dB; it is shown in (1).

$$SDR = 10\log_{10}\left(\frac{Signal_{original}}{Signal_{estimated}}\right) \tag{1}$$

The formula denotes the higher the SDR, the higher the $signal_{original}$, the lower the $signal_{estimated}$. In other words, higher SDR value represents estimated signal is more incline toward original signal which it considered as more satisfied.

2. PROPOSED METHOD

2.1. General framework of α-divergence

Nonnegative matrix factorization (NMF) defines finds two non-negative matrices which are $W \in R^{n\times r}$ and $H \in R^{r\times m}$, which also can be concluded as $V \approx WH$. There is bunch of famous cost function has been deployed to auxiliary NMF such as Dual Kullback-Leibler I (DKLI) divergence, Squared Hellinger (SH) divergence, Kullback-Leibler I (KLI) divergence and Pearson divergence. The divergences that mentioned above are derived from a general framework which is α-divergence [24], [25]. The framework of α-divergence is shown as (2),

$$d_\alpha(V|\Lambda) = \begin{cases} \frac{1}{\alpha(\alpha-1)}[V^\alpha(\Lambda)^{1-\alpha} - \alpha V + (\alpha-1)(\Lambda)] & \alpha \in \Re\ \{0,2\} \\ \Lambda\log\left(\frac{\Lambda}{V}\right) + V - \Lambda & \alpha = 0 \\ -4\left(\sqrt{V\Lambda} - \frac{V+\Lambda}{2}\right) & \alpha = 0.5 \\ V\log\left(\frac{V}{\Lambda}\right) - V + \Lambda & \alpha = 1 \\ \frac{1}{2}\left(\frac{V^2}{\Lambda} + \Lambda - 2V\right) & \alpha = 2 \end{cases} \tag{2}$$

where $d_\beta(y|x)$ is scalar cost function and $\alpha \in \Re\ \{0,2\}$. The DKLI divergence, SH divergence, KLI divergence and Pearson divergence represent α= 0, 0.5, 1 and 2 respectively [24]. In NMF presentation, all of the divergence will turn into (3) to (6) via substitution of different α value in α-divergence framework.

$$C_{\mathrm{DKLI}} = \sum_i\sum_j \Lambda_{ij}\log\left(\frac{\Lambda_{ij}}{V_{ij}}\right) + V_{ij} - \Lambda_{ij} \tag{3}$$

$$C_{\mathrm{DKLI}} = \sum_i\sum_j -4\left(\sqrt{V_{ij}\Lambda_{ij}} - \frac{V_{ij}+\Lambda_{ij}}{2}\right) \tag{4}$$

$$C_{\mathrm{DKLI}} = \sum_i\sum_j V_{ij}\log\left(\frac{V_{ij}}{\Lambda_{ij}}\right) - V_{ij} + \Lambda_{ij} \tag{5}$$

$$C_{\mathrm{DKLI}} = \sum_i\sum_j \frac{1}{2}\left(\frac{V_{ij}^2}{\Lambda_{ij}} + \Lambda_{ij} - 2V_{ij}\right) \tag{6}$$

The BSS in this paper is classified into single channel source separation (SCSS). In time domain, model of SCSS is (7).

$$V(t) = \sum_{j=1}^{J}\Lambda_j(t) + e(t) \tag{7}$$

It is then change into time-frequency domain via short time Fourier transform (STFT),

$$V(t) = \sum_{j=1}^{J} \Lambda_{j.f,n} + e_{f,n} \tag{8}$$

where $j = 1,2,3, \dots, J$ denotes amount of source, *e(t)* denotes additional interference, $f = 1,2,3, \dots, F$ denotes frequency bin and $n = 1,2,3, \dots, N$ denotes time frame index.

$$\left|X_j\right|^2 = \sum_{\tau=0}^{\tau_{max}} \sum_{\phi=0}^{\phi_{max}} \overset{\downarrow\phi}{W_j^{\tau}} \overset{\rightarrow\tau}{H_j^{\phi}} \tag{9}$$

The matrix W shows the τ^{th} slice spectral basis and H shows the ϕ^{th} slice of temporal code for each spectral basis element. Arrow of $\overset{\downarrow\phi}{W_j^{\tau}}$ shows shifting each element by ϕ row down and arrow of $\overset{\rightarrow\tau}{H_j^{\phi}}$ shows shifting each element by τ column right [21].

2.2. Multiplicative update rules without sparseness constraints

Due to the one disadvantage of the basic NMF formulation is its inconsistency to manipulate the amount of dependence among the learned dictionary atoms, MU rules present the convenient method to enforce dependence among atoms. In fact, this means that repeated iteration of the update rules is guaranteed to converge to a locally optimal matrix factorization [26]. Therefore, in this paper, we deployed MU rules on the NMF2D α-divergence by adding multiplicative gradient descent method initially with positive learning rate in order to grant the sparse cost function to reach the minimum [27]. Now, we elaborate the α-divergence as defined in (2),

$$D_{\alpha}\left(|V|^2 \middle| \tilde{\Lambda}\right) = \sum_{f,n} \left[\frac{1}{\alpha(\alpha-1)} \left[|V|_{f,n}^{2\ \alpha} \left(\tilde{\Lambda}_{f,n}\right)^{1-\alpha} - \alpha |V|_{f,n}^2 + (\alpha-1)\left(\tilde{\Lambda}_{f,n}\right)\right]\right] \tag{10}$$

where $\tilde{V} = \sum_{j,\tau,\phi} \overset{\downarrow\phi}{W_j^{\tau}} \overset{\rightarrow\tau}{H_j^{\phi}}$ with $\widetilde{W}_{f,j}^{\tau} = \frac{W_{f,j}^{\tau}}{\sqrt{\sum_{\tau,f}\left(W_{f,j}^{\tau}\right)^2}}$ in $f = 1, \dots, F$, $n = 1, \dots, N$ and parameter λ is the sparsity constraint. Then, the derivatives of (10) are given by (11) and (12):

$$\begin{aligned} \frac{\delta D_{\alpha}}{\delta W_{f',j'}^{\tau'}} &= \frac{\delta}{\delta W_{f',j'}^{\tau'}} \left(\sum_{f,n} \left(\frac{1}{\alpha(\alpha-1)} \left[|V|_{f,n}^{2\ \alpha} \left(\tilde{\Lambda}_{f,n}\right)^{1-\alpha} - \alpha|V|_{f,n}^2 + (\alpha-1)\left(\tilde{\Lambda}_{f,n}\right)\right]\right)\right) \\ &= \sum_{\phi,n} \left(\frac{\left(\tilde{\Lambda}_{f,n}\right)^{\alpha} - |V|_{f,n}^{2\ \alpha}}{\alpha\left(\left(\tilde{\Lambda}_{f,n}\right)^{\alpha}\right)}\right) H_{j',n-\tau'}^{\phi} \end{aligned} \tag{11}$$

$$\begin{aligned} \frac{\delta D_{\alpha}}{\delta H_{j',n'}^{\phi'}} &= \frac{\delta}{\delta H_{j',n'}^{\phi'}} \left(\sum_{f,n} \left(\frac{1}{\alpha(\alpha-1)} \left[|V|_{f,n}^{2\ \alpha} \left(\tilde{\Lambda}_{f,n}\right)^{1-\alpha} - \alpha|V|_{f,n}^2 + (\alpha-1)\left(\tilde{\Lambda}_{f,n}\right)\right]\right)\right) \\ &= \sum_{f,\phi} \widetilde{W}_{f-\phi',j'}^{\tau} \left(\frac{\left(\tilde{\Lambda}_{f,n}\right)^{\alpha} - |V|_{f,n}^{2\ \alpha}}{\alpha\left(\left(\tilde{\Lambda}_{f,n}\right)^{\alpha}\right)}\right) \end{aligned} \tag{12}$$

For gradient descent method [19], it was shown as (13).

$$W_{f',j'}^{\tau'} \leftarrow \widetilde{W}_{f',j'}^{\tau'} - \eta_W \frac{\delta D_{\alpha}}{\delta W_{f',j'}^{\tau'}} \tag{13}$$

$$H_{j',n'}^{\phi'} \leftarrow \widetilde{H}_{j',n'}^{\phi'} - \eta_H \frac{\delta D_{\alpha}}{\delta H_{j',n'}^{\phi'}} \tag{14}$$

Therefore, after substituted into (13) and (14), multiplicative rules become (15) and (16).

$$W^{\tau} \leftarrow \widetilde{W^{\tau}} \cdot \frac{\sum_{\phi} \left[\begin{pmatrix} \uparrow \phi \\ |V|^2 \end{pmatrix}^{\alpha} \right] \overset{\rightarrow \tau}{H^{\phi}}^{T}}{\sum_{\phi} \begin{pmatrix} \uparrow \phi \\ \widetilde{\Lambda} \end{pmatrix}^{\alpha} \overset{\rightarrow \tau}{H^{\phi}}^{T}} \tag{15}$$

$$H^{\phi} \leftarrow \widetilde{H^{\phi}} \cdot \frac{\sum_{\tau} \overset{\downarrow \phi}{W^{\tau}}^{T} \left[\begin{pmatrix} \uparrow \phi \\ |V|^2 \end{pmatrix}^{\alpha} \right]}{\sum_{\tau} \overset{\downarrow \phi}{W^{\tau}}^{T} \begin{pmatrix} \uparrow \phi \\ \widetilde{\Lambda} \end{pmatrix}^{\alpha}} \tag{16}$$

2.3. Multiplicative update rules with sparseness constraints

Sparseness constraints [26] were deployed to diminish the ambiguity and provide uniqueness to the solution. In addition, in the way of sparseness constraints representation, much of the data of encodes process using few 'active' components, which grants the encoding easy to interpret. In other words, matrix considered dense if most of the elements are nonzero and vice versa [4], [28]. Now, we elaborate the α-divergence as defined in (1) by addition of sparseness constraints in order to diminish the cost function as (17),

$$\frac{\delta D_{\alpha}}{\delta W^{\tau'}_{f',j'}} = \frac{\delta}{\delta W^{\tau'}_{f',j'}} \left(\sum_{f,n} \left(\frac{1}{\alpha(\alpha-1)} \left[|V|^{2\ \alpha}_{f,n} \left(\widetilde{\Lambda}_{f,n}\right)^{1-\alpha} - \alpha |V|^{2}_{f,n} + (\alpha-1)\left(\widetilde{\Lambda}_{f,n}\right) \right] \right) + \lambda f(H) \right) \tag{17}$$

where $\tilde{V} = \sum_{j,\tau,\phi} \overset{\downarrow \phi \rightarrow \tau}{W^{\tau}_{j} H^{\phi}_{j}}$ with $\widetilde{W}^{\tau}_{f,j} = \frac{W^{\tau}_{f,j}}{\sqrt{\sum_{\tau,f} \left(W^{\tau}_{f,j}\right)^2}}$ in $f = 1, \dots, F,\ n = 1, \dots, N$ and parameter λ is the sparsity constraint. Then, the derivatives of (9) are given by (18) and (19).

$$\begin{aligned} \frac{\delta D_{\alpha}}{\delta W^{\tau'}_{f',j'}} &= \frac{\delta}{\delta W^{\tau'}_{f',j'}} \left(\sum_{f,n} \left(\frac{1}{\alpha(\alpha-1)} \left[|V|^{2\ \alpha}_{f,n} \left(\widetilde{\Lambda}_{f,n}\right)^{1-\alpha} - \alpha |V|^{2}_{f,n} + (\alpha-1)\left(\widetilde{\Lambda}_{f,n}\right) \right] \right) + \lambda f(H) \right) \\ &= \sum_{\phi,n} \left(\frac{\left(\widetilde{\Lambda}_{f,n}\right)^{\alpha} - |V|^{2\ \alpha}_{f,n}}{\alpha\left(\left(\widetilde{\Lambda}_{f,n}\right)^{\alpha}\right)} \right) H^{\phi}_{j',n-\tau'} \end{aligned} \tag{18}$$

$$\begin{aligned} \frac{\delta D_{\alpha}}{\delta W^{\tau'}_{f',j'}} &= \frac{\delta}{\delta W^{\tau'}_{f',j'}} \left(\sum_{f,n} \left(\frac{1}{\alpha(\alpha-1)} \left[|V|^{2\ \alpha}_{f,n} \left(\widetilde{\Lambda}_{f,n}\right)^{1-\alpha} - \alpha |V|^{2}_{f,n} + (\alpha-1)\left(\widetilde{\Lambda}_{f,n}\right) \right] \right) + \lambda f(H) \right) \\ &= \sum_{f,\phi} \widetilde{W}^{\tau}_{f-\phi',j'} \left(\frac{\left(\widetilde{\Lambda}_{f,n}\right)^{\alpha} - |V|^{2\ \alpha}_{f,n}}{\alpha\left(\left(\widetilde{\Lambda}_{f,n}\right)^{\alpha}\right)} \right) + \lambda \frac{\delta f(H)}{\delta H^{\phi'}_{j',n'}} \end{aligned} \tag{19}$$

For gradient descent method [19], it was shown as (21).

$$W^{\tau'}_{f',j'} \leftarrow \widetilde{W}^{\tau'}_{f',j'} - \eta_W \frac{\delta D_{\alpha}}{\delta W^{\tau'}_{f',j'}} \tag{20}$$

$$H^{\phi'}_{j',n'} \leftarrow \widetilde{H}^{\phi'}_{j',n'} - \eta_H \frac{\delta D_{\alpha}}{\delta H^{\phi'}_{j',n'}} \tag{21}$$

Therefore, after substituted into (20) and (21), multiplicative rules become (22) and (23).

$$W^{\tau} \leftarrow \widetilde{W^{\tau}} \cdot \frac{\sum_{\phi} \left[\begin{pmatrix} \uparrow \phi \\ |V|^2 \end{pmatrix}^{\alpha} \right] \overset{\rightarrow \tau}{H^{\phi}}^{T}}{\sum_{\phi} \begin{pmatrix} \uparrow \phi \\ \widetilde{\Lambda} \end{pmatrix}^{\alpha} \overset{\rightarrow \tau}{H^{\phi}}^{T}} \tag{22}$$

$$H^{\phi} \leftarrow \widetilde{H^{\phi}} \cdot \frac{\sum_{\tau} \overset{\downarrow \phi}{W^{\tau}}^{T} \left[\begin{pmatrix} \uparrow \phi \\ |V|^2 \end{pmatrix}^{\alpha} \right]}{\sum_{\tau} \overset{\downarrow \phi}{W^{\tau}}^{T} \left[\begin{pmatrix} \uparrow \phi \\ \widetilde{\Lambda} \end{pmatrix}^{\alpha} \left(1 + \left(\lambda \frac{\delta f(H)}{\delta H^{\phi}} \right) (\alpha) \right) \right]} \tag{23}$$

3. RESULTS AND DISCUSSION

All simulations and analyses are run using MATLAB platform. Mixture of heart and lung sound signal are sampled at 44.1 kHz. The convolutive components in time and frequency are selected to be $\tau\_max=3$ and $\phi\_max=31$ for every case and the performance of results will be processed through comparison of estimated to original audio signal in term of signal-to-distortion ratio (SDR).

3.1. Condition without sparseness constraints

We implemented α with the step size of 0.1 which is starting from 0 until 2. It ought to consist of general framework of α divergence family which is DKLI divergence, SH divergence, KLI divergence and Pearson divergence. According to Table 1, the peak point is at α=1.5 which contains SDR=17.8396 dB. From Figure 1, it reveals the SDR increased dramatically at α=0.1, it next fluctuated within range of SDR=17 to SDR=18 after α=0.9.

Table 1. The influences of the α with step size of 0.1 to SDR's average

Separated estimated sound / α	SDR's average (dB)
0	1.9341
0.1	13.9447
0.2	15.3062
0.3	15.9714
0.4	16.8237
0.5	16.1725
0.6	15.5676
0.7	15.5739
0.8	15.4812
0.9	17.5545
1.0	17.7698
1.1	17.7158
1.2	17.6698
1.3	17.6936
1.4	17.7668
1.5	17.8396
1.6	17.6989
1.7	17.7861
1.8	17.8270
1.9	17.8147
2.0	17.1793

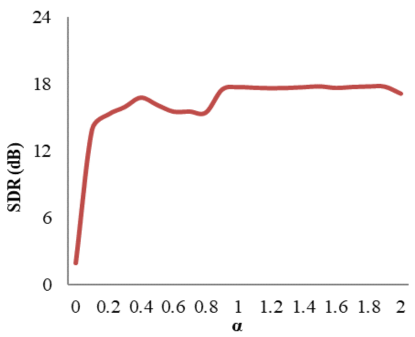
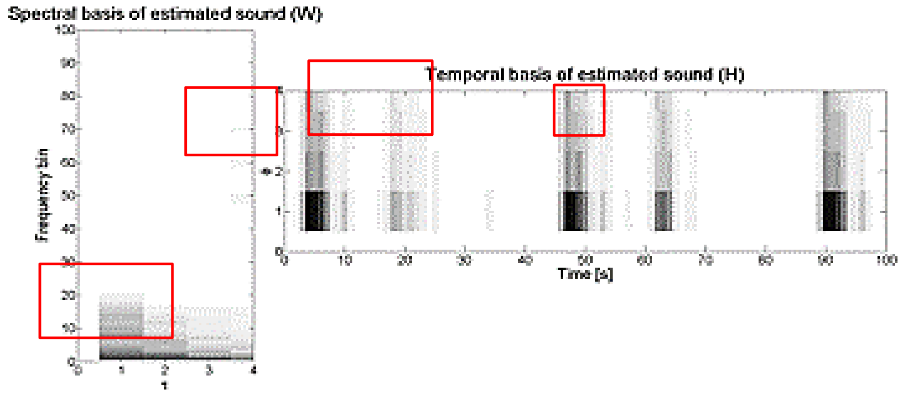


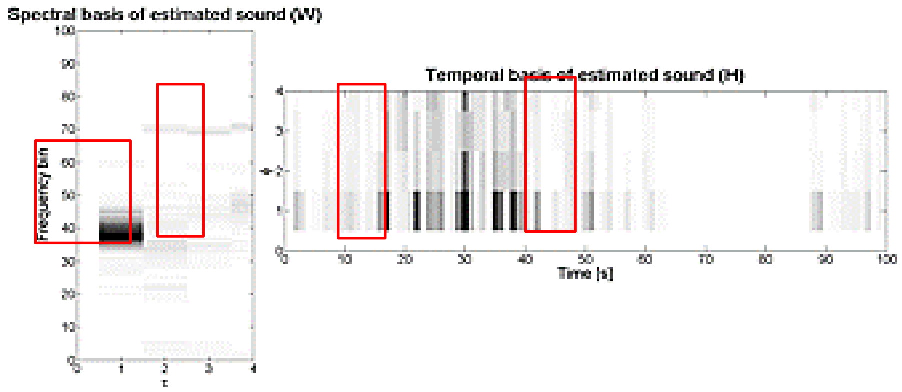
Figure 1. SDR (dB) against α

3.2. Condition with sparseness constraints

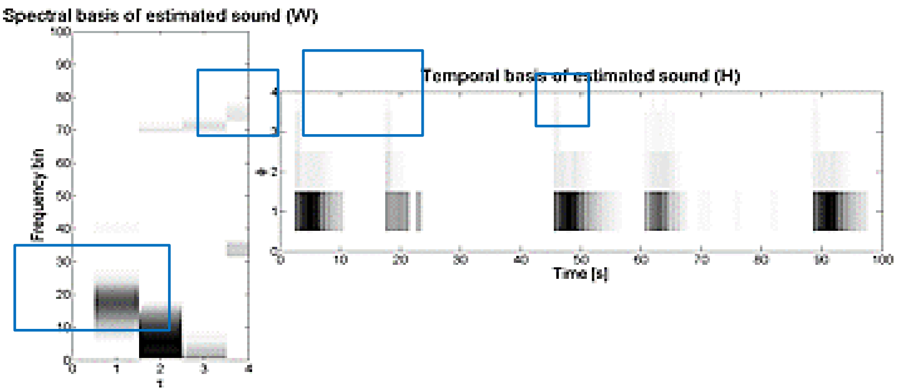
Figure 2 reveals mixture of heart and lung sound after decomposition with and without sparseness constraints in term of matrix W (spectral basis) and matrix H (temporal basis). Combination of matrix W and matrix H will be constructed into V. Matrix W and H is produced under NMF2D which been discussed previously. Figures 2(a) and 2(b) show the huge portion of light gray color (marked as red boxes) compared to Figures 2(c) and 2(d) (marked as blue boxes). This is due to the ambiguity of signal is truncated when the auxiliary constraints are added. Thus, this remarks that when sparseness constraints are added, the estimated signal is more incline to the original signal which cost function is reduced almost to zero, or in other words, the difference between estimated and original signal been minimize.



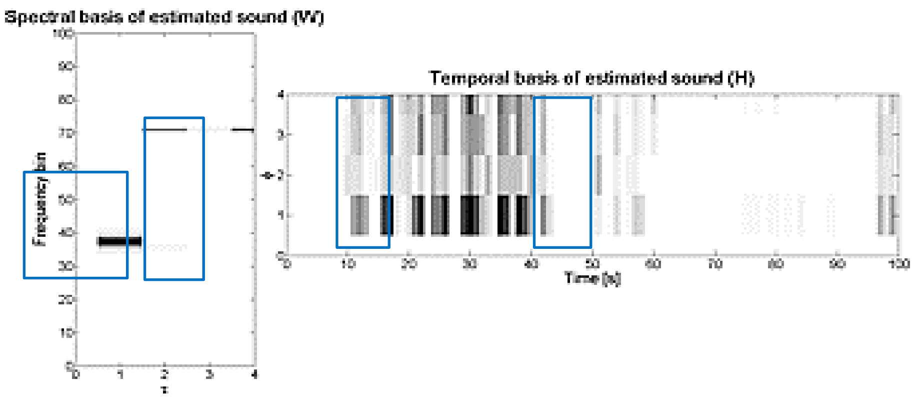
(a)



(b)



(c)



(d)

Figure 2. The estimated W and H of heart sound without sparseness constraints for (a) heart sound and (b) lung sound, and with sparseness constraints for (c) heart sound and (d) lung sound

4. CONCLUSION

In conclusion, we verify that $\alpha=1.5$ as the optimal value in family of α-divergence in term of SDR. At the meanwhile, we deploy sparseness constraints in NMF2D algorithm and found that SDR value even higher than NMF2D without sparseness constraint which is at $\lambda=4.5$. Thus, $\lambda=4.5$ is considered as best value to eliminate the ambiguity and grant estimated signal almost near to original signal among the other λ value. Hence, its performance such as SDR, speed of decomposition iteration as well as weight of matrix has been improved. In the future, we believe that these constraints will become helpful in various application addressed by NMF2D beyond audio source separation.

ACKNOWLEDGEMENTS

The author would like to express our gratitude to honorable university, Universiti Teknikal Malaysia Melaka (UTeM). Special appreciation especially to Centre of Research and Innovation Management (CRIM), Machine Learning and Signal Processing (MLSP) Research Group, and Faculty of Electronic and Computer Engineering, UTeM.

REFERENCES

[1] P. Paatero and U. Tapper, "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics*, vol. 5, no. 2, pp. 111–126, Jun. 1994, doi: 10.1002/env.3170050203.

[2] J.-T. Chien, *Source separation and machine learning*. Elsevier, 2019.

[3] A. Darsono, B. Gao, W. L. Woo, and S. S. Dlay, "Nonlinear single channel source separation," in *2010 7th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP 2010)*, Jul. 2010, pp. 507–511, doi: 10.1109/CSNDSP16145.2010.5580373.

[4] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *Journal of Machine Learning Research*, vol. 5, pp. 1457–1469, 2004.

[5] J. Eggert and E. Korner, "Sparse coding and NMF," in *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*, 2004, vol. 4, pp. 2529–2533, doi: 10.1109/IJCNN.2004.1381036.

[6] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999, doi: 10.1038/44565.

[7] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Advances in Neural Information Processing Systems*, 2001, pp. 556–562.

[8] R. Ranny, D. P. Lestari, T. Latifah Erawati Rajab, and I. S. Suwardi, "Separation of overlapping sound using nonnegative matrix factorization," in *2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, Dec. 2019, pp. 424–429, doi: 10.1109/ISRITI48646.2019.9034580.

[9] D. Donoho and V. Stodden, "When does non-negative matrix factorization give a correct decomposition into parts?," *MIT Press*, vol. 16, pp. 1141–1148, 2004.

[10] S. Sra and I. Dhillon, "Generalized nonnegative matrix approximations with Bregman divergences," in *Advances in Neural Information Processing Systems 18 (NIPS 2005)*, 2006, vol. 18, pp. 283–290.

[11] C.-J. Lin, "Projected gradient methods for nonnegative matrix factorization," *Neural Computation*, vol. 19, no. 10, pp. 2756–2779, Oct. 2007, doi: 10.1162/neco.2007.19.10.2756.

[12] Chih-Jen Lin, "On the convergence of multiplicative update algorithms for nonnegative matrix factorization," *IEEE Transactions on Neural Networks*, vol. 18, no. 6, pp. 1589–1596, Nov. 2007, doi: 10.1109/TNN.2007.895831.

[13] H. Pasterkamp, S. S. Kraman, and G. R. Wodicka, "Respiratory sounds," *American Journal of Respiratory and Critical Care Medicine*, vol. 156, no. 3, pp. 974–987, Sep. 1997, doi: 10.1164/ajrccm.156.3.9701115.

[14] X. Lin and P. C. Boutros, "Optimization and expansion of non-negative matrix factorization," *BMC Bioinformatics*, vol. 21, no. 1, p. 7, Dec. 2020, doi: 10.1186/s12859-019-3312-5.

[15] N. Yu, J.-X. Liu, Y.-L. Gao, C.-H. Zheng, J. Wang, and M.-J. Wu, "Graph regularized robust non-negative matrix factorization for clustering and selecting differentially expressed genes," in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Nov. 2017, pp. 1752–1756, doi: 10.1109/BIBM.2017.8217925.

[16] V. Leplat, N. Gillis, and A. M. S. Ang, "Blind audio source separation with minimum-volume beta-divergence NMF," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3400–3410, 2020, doi: 10.1109/TSP.2020.2991801.

[17] M. Hameed, M. M. Ibrahim, N. A. Manap, and A. A. Mohammed, "An enhanced lossless compression with cryptography hybrid mechanism for ECG biomedical signal monitoring," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 3, pp. 3235–3243, Jun. 2020, doi: 10.11591/ijece.v10i3.pp3235-3243.

[18] V. J. Naveen, K. M. Krishna, and K. R. Rajeswari, "Noise reduction in ECG signals for bio-telemetry," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 1, pp. 505–511, Feb. 2019, doi: 10.11591/ijece.v9i1.pp505-511.

[19] N. A. Nayan and H. A. Hamid, "Evaluation of patient electrocardiogram datasets using signal quality indexing," *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 2, pp. 519–526, Jun. 2019, doi: 10.11591/eei.v8i2.1289.

[20] Y. Teng, J. Yan, and Y. Kang, "Iterative non-negative deconvolution algorithms with alpha-divergence," in *2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics*, Aug. 2015, pp. 393–398, doi: 10.1109/IHMSC.2015.224.

[21] M. Morup and M. N. Schmidt, "Sparse nonnegative matrix factor 2-D deconvolution," Technical Report, Technical University of Denmark, Denmark, 2006.

[22] A. Cichocki, H. Lee, Y.-D. Kim, and S. Choi, "Non-negative matrix factorization with α-divergence," *Pattern Recognition Letters*, vol. 29, no. 9, pp. 1433–1440, Jul. 2008, doi: 10.1016/j.patrec.2008.02.016.

[23] E. Vincent, R. Gribonval, and C. Fevotte, "Performance measurement in blind audio source separation," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 14, no. 4, pp. 1462–1469, Jul. 2006, doi: 10.1109/TSA.2005.858005.

[24] S. Ferdowsi, V. Abolghasemi, and S. Sanei, "A comparative study of α−divergence based NMF techniques for FMRI analysis," in *19th European Signal Processing Conference (EUSIPCO 2011)*, 2011, pp. 71–75.

[25] A. Cichocki, S. Cruces, and S. Amari, "Generalized alpha-beta divergences and their application to robust nonnegative matrix factorization," *Entropy*, vol. 13, no. 1, pp. 134–170, Jan. 2011, doi: 10.3390/e13010134.

[26] S. K. Tjoa and K. J. R. Liu, "Multiplicative update rules for nonnegative matrix factorization with co-occurrence constraints," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, Mar. 2010, pp. 449–452, doi: 10.1109/ICASSP.2010.5495734.

[27] L. Sui, X. Zhang, J. Huang, G. Zhao, and Y. Yang, "Speech enhancement based on sparse nonnegative matrix factorization with priors," in *2012 International Conference on Systems and Informatics (ICSAI2012)*, May 2012, pp. 274–278, doi: 10.1109/ICSAI.2012.6223615.

[28] A. M. Darsono, N. Z. Haron, A. S. Jaafar, and M. I. Ahmad, "β-divergence two-dimensional sparse nonnegative matrix factorization for audio source separation," in *2013 IEEE Conference on Wireless Sensor (ICWISE)*, Dec. 2013, pp. 119–123, doi: 10.1109/ICWISE.2013.6728792.

BIOGRAPHIES OF AUTHORS

Abd Majid Darsono received the bachelor's degree in communication and computer engineering from Universiti Kebangsaan Malaysia in 2002 and the M.Sc. degree in electronic communications and computer engineering from The University of Nottingham, United Kingdom in 2005. In 2012, he received the Ph.D. degree in signal processing from University of Newcastle upon Tyne, United Kingdom. He joined Universiti Teknikal Malaysia Melaka (UTeM) since 2005 as a lecturer. He is currently an associate professor in Faculty of Electronics and Computer Engineering in UTeM. His research experience and interests include statistical signal processing, and speech and image processing. He can be contacted at abdmajid@utem.edu.my.

Toh Cheng Chuan received his bachelor's degree in Electronic and Computer Engineering from Technical University of Malaysia Malacca in 2015 and obtained his master's degree in the same university in 2018, majoring in signal processing. His research interests include blind source separation and biomedical signal processing. He can be contacted at tohchengchuan@gmail.com.

Nurulfajar Abd Manap serves as a faculty member of the Faculty of Electronics and Computer Engineering at the Universiti Teknikal Malaysia Melaka. He earned a Master of Electrical Engineering in image processing (2002, Universiti Teknologi Malaysia) and a Bachelor of Electrical Engineering (2000, Universiti Teknologi Malaysia). He obtained his Ph.D. (image and video processing) at the University of Strathclyde, Glasgow in 2012. His research interests are 3D image processing, stereo vision, virtual and augmented reality, and video processing. He can be contacted at nurulfajar@utem.edu.my.

Nik Mohd Zarifie Hashim received the B.Eng. and M.Eng. degrees in Electrical and Electronics Engineering from University of Fukui, Japan in 2006 and 2008. In 2007, he joined Universiti Teknikal Malaysia Melaka as a tutor. In 2008, he became lecturer in the same university. Since 2013, he has been a senior lecturer there. During 2016 to 2019, he was a Ph.D. student at Graduate School of Information Science, Nagoya University, Japan. He is member of IEEE, graduate engineer of Board of Engineering Malaysia (BEM), graduate technologists of Malaysia Board of Technologists (MBOT), and graduate member of The Institution of Engineers, Malaysia (IEM). His research focuses on person re-identification and 3D object pose estimation. He can be contacted at nikzarifie@utem.edu.my.